DOCKET NO. 2000-0152US1

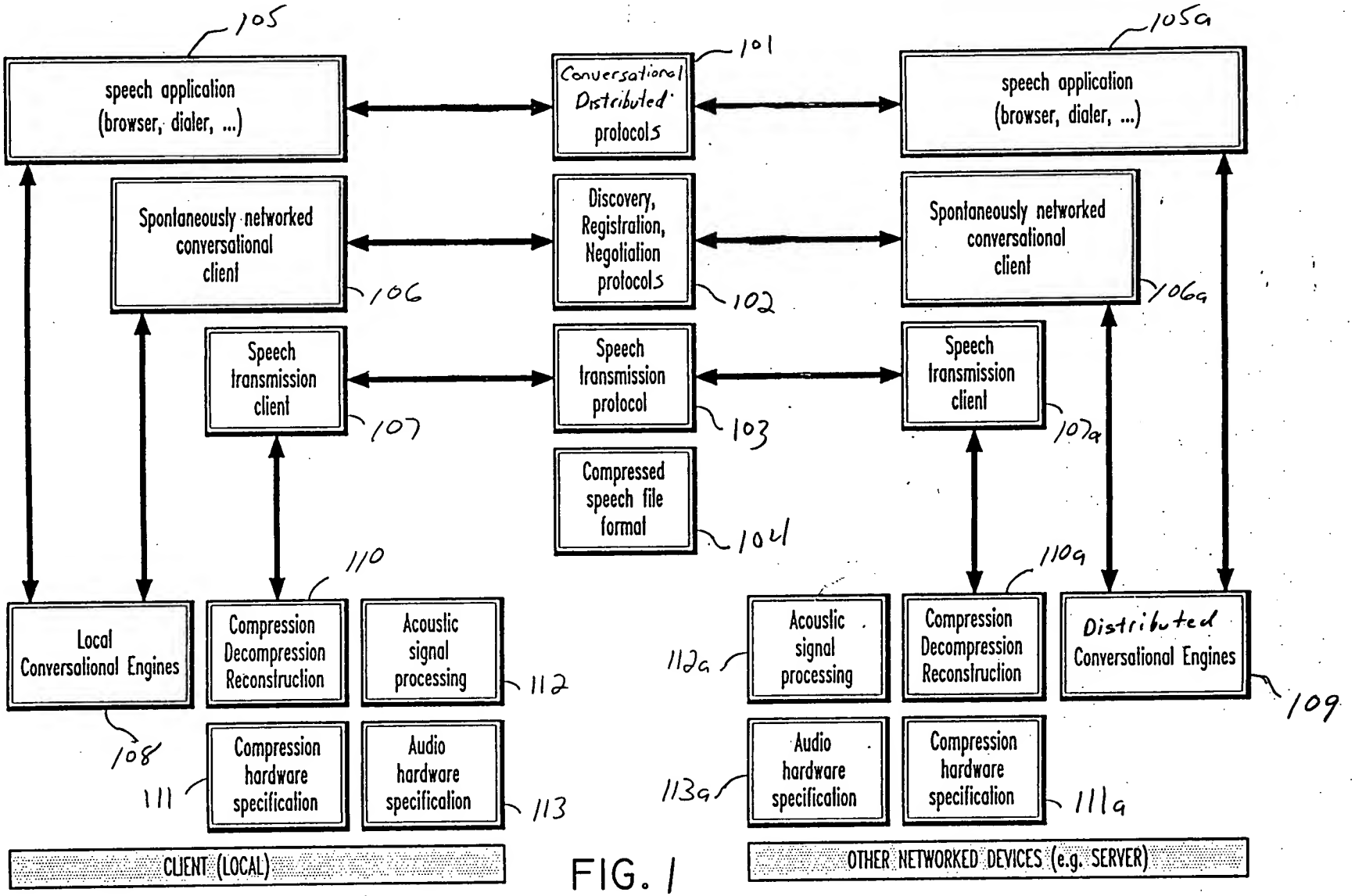

FIG. 1

S. Maes 1/18 2000-0152US1 (P) 8726-376

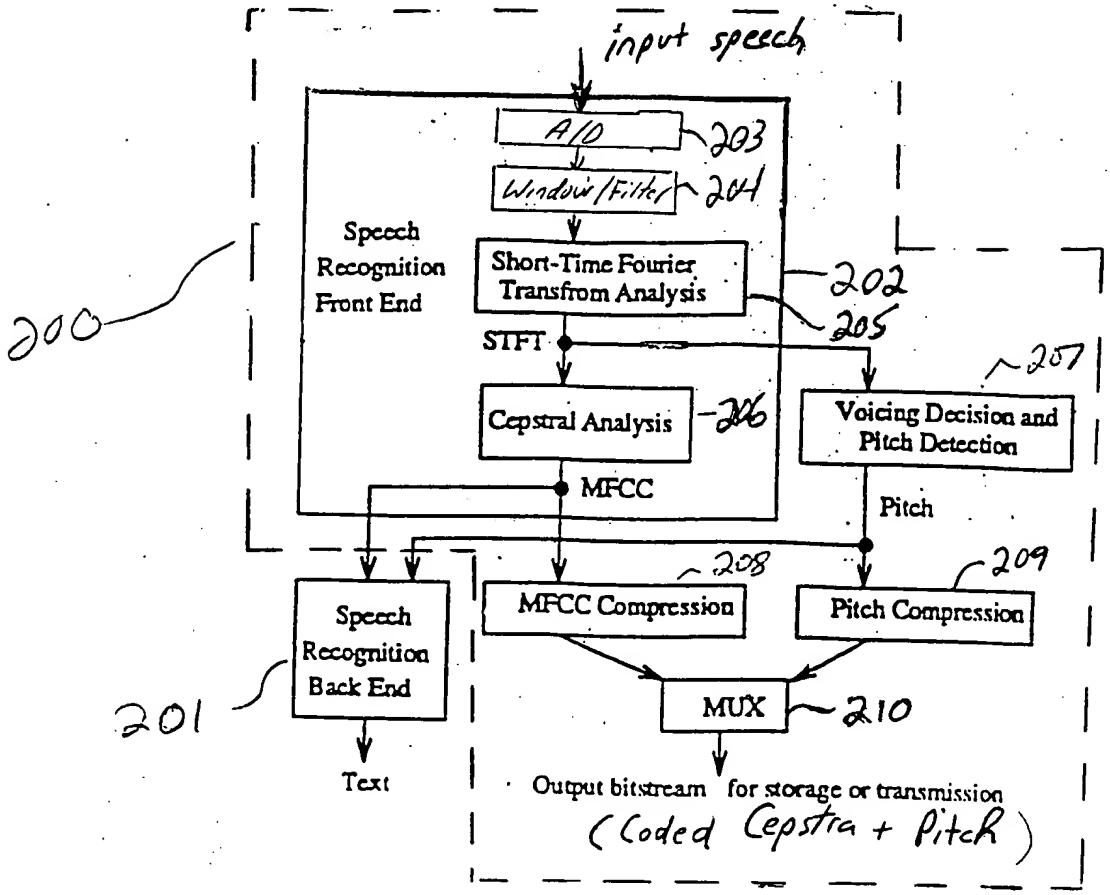
input speech
A/D
203
Window/Filter
204
Speech Recognition Front End
Short-Time Fourier Transfrom Analysis
202
205
STFT
207
Cepstral Analysis
206
Voicing Decision and Pitch Detection
MFCC
Pitch
208
209
Speech Recognition Back End
MFCC Compression
Pitch Compression
MUX
210
200
201
Text
Output bitstream for storage or transmission
(Coded Cepstra + Pitch)

Fig. 2a

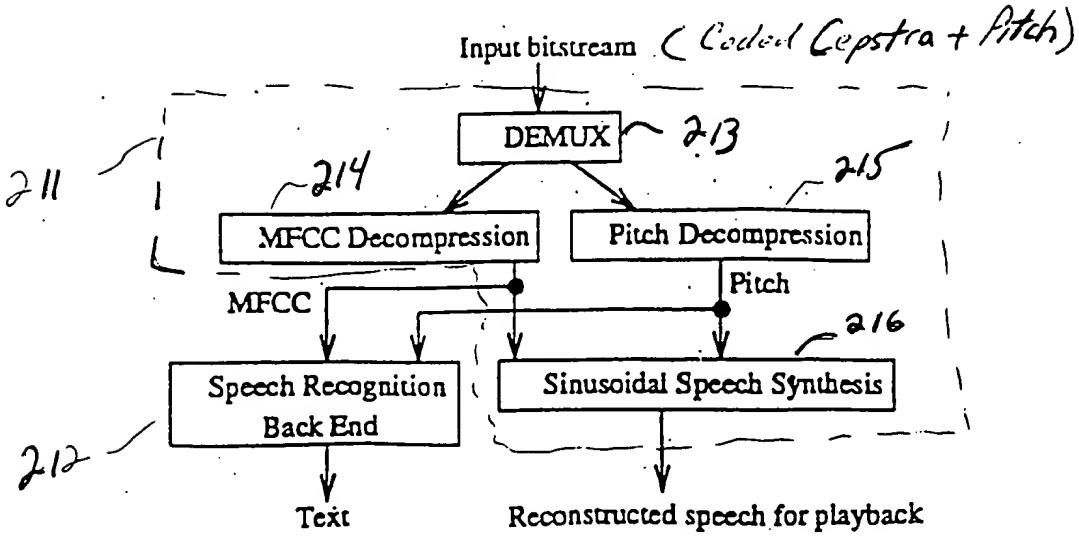
Input bitstream (Coded Cepstra + Pitch)
DEMUX
213
211
214
215
MFCC Decompression
Pitch Decompression
Pitch
MFCC
216
Speech Recognition Back End
Sinusoidal Speech Synthesis
212
Text
Reconstructed speech for playback

Fig 2b

RecoVC File Format

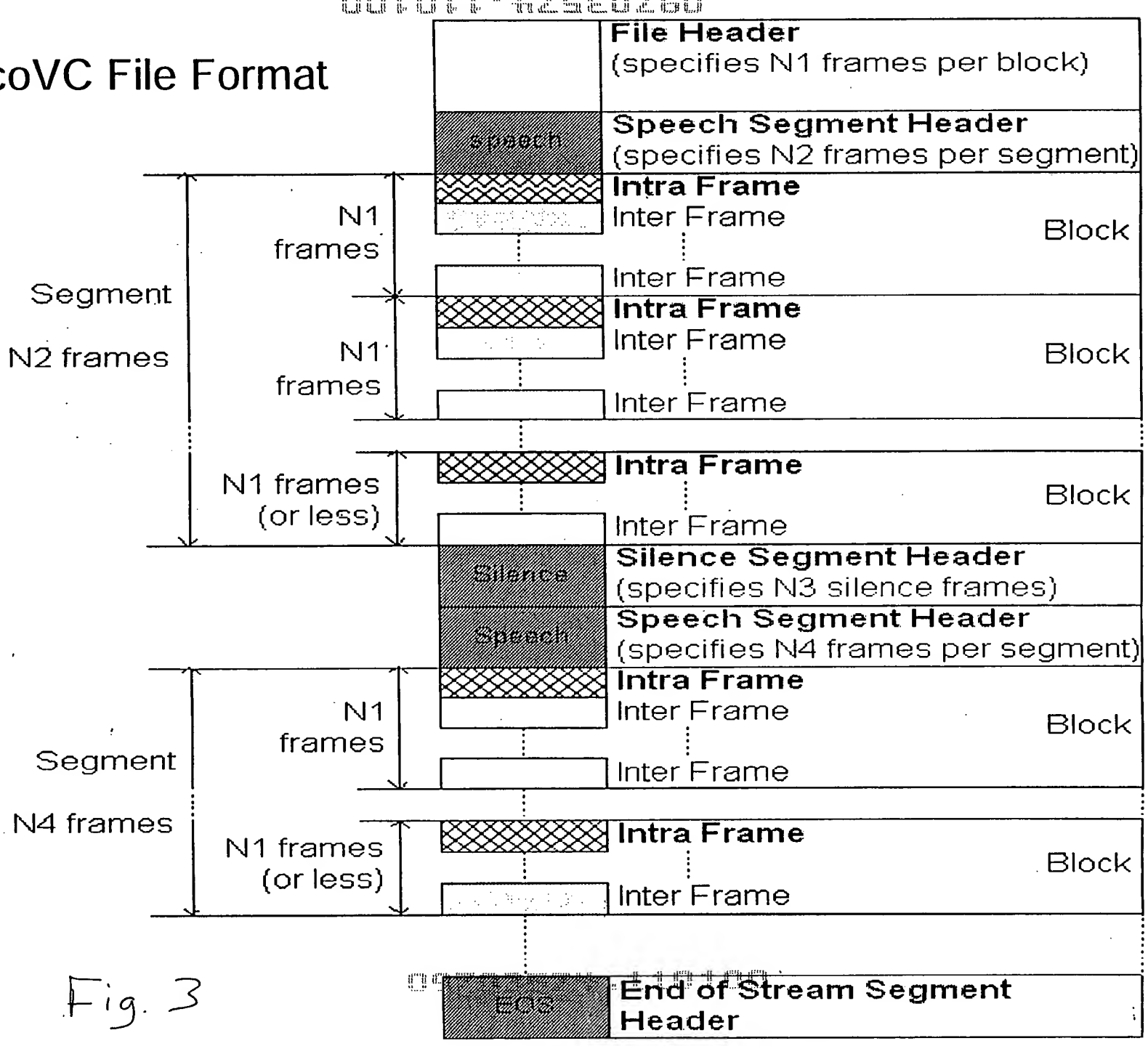

Fig. 3

3/18

YOR9-2000-0154US1 (6228-376)

00101 "4259268

4/18
YOR9-2002-0154 US1 (8728-376)

RecoVC File Header

Field Name	M/O	Number of bytes	Value(s)	Description
Recovc magic number	M	4 bytes	0x2E727663	A fixed value which equals ".rvc" in ASCII representation. Used to verify that this is indeed a Recovc file.
Header length	M	2 byte	13-(2^16-1)	Total number of bytes in this header
Version Major	M	1 byte	0-255	File format major version number
Version Minor	M	1 byte	0-255	File format minor version number
Frame duration	M	1 byte	0x00	In index into the frame duration table below. Currently, only 10 msec frames are supported.
Frames per block	M	1 byte	1-255	Each block can contain one intra-frame only, or one intra-frame with up to 244 additional inter-frames.
Sampling rate	M	1 byte	0x00, 0x01 or 0x02	Index into sampling rate table below.
Cepstra dimension	M	1 byte	0x00 or 0x01	Index into Cepstra vector dimension / type below
Language	M	1 byte	0x00	Index into languate table below
Profile	M	1 byte	0x00 or 0x01	Index into profile table below
Total frames	O	4 bytes	0-(2^32-2) or 0xFFFFFFFF	Total number of audio frames in this file, INCLUDING silence frames.
Total speech frames	O	4 bytes	0-(2^32-2) or 0xFFFFFFFF	Total number of audio frames in this file, EXCLUDING silence frames.
< Ancillary Data >	O			Reserved for future usage

Fig 4

Frame Duration Table

Code	Duration (msec)
0x00	10
0x01- 0xFF	Reserved

Sampling Rate Table

Code	Input Sampling Rate [Hz]
0x00	8000
0x01	11025 (or 11000)
0x02	16000*
0x03	22050 (or 22000)
0x04- 0xFF	Reserved

Cepstra Vector Dimension / type Table

Code	Value
0x00	13
0x01	24
0x02- 0xFF	Reserved

Language Table

Code	Description
0x00	English
0x01- 0xFF	Reserved

Fig. 5

5/18
Vol9-2000-0154 US1 (8728-376)

6/18
Y029-2000-054US1(6728-376)

Profile Table

Code	Description
0x00	Recognition only - Only Cepstra is available in this file
0x01	Reconstruction enabled - Cepstra and pitch is available in the file
0x02- 0xFF	Reserved

Speech and Silence Segment Headers (5 Byte)

Field Name	M/O	Number of bytes	Value(s)	Description
Segment type	M	1 byte	0x00 or 0x01	Index into segment type table below (speech or silence)
Number of frames	M	4 bytes	0-(2^31-2) or 0xFFFFFFFF	The total number of frames in this segment.

End of Stream and Data Segment Headers

Field Name	M/O	Number of bytes	Value(s)	Description
Segment type	M	1 byte	0x02 or 0x03-0xFF	Index into segment type table below (end of stream or data)
Segment length	M	4 bytes	5-(2^31-1)	The total number of bytes in this segment
< Ancillary Data >	O	.	.	Place optional data here

Fig 6

00F0FT "62E0260

7/18
YOR9-2000-0154US1 (8728-376)

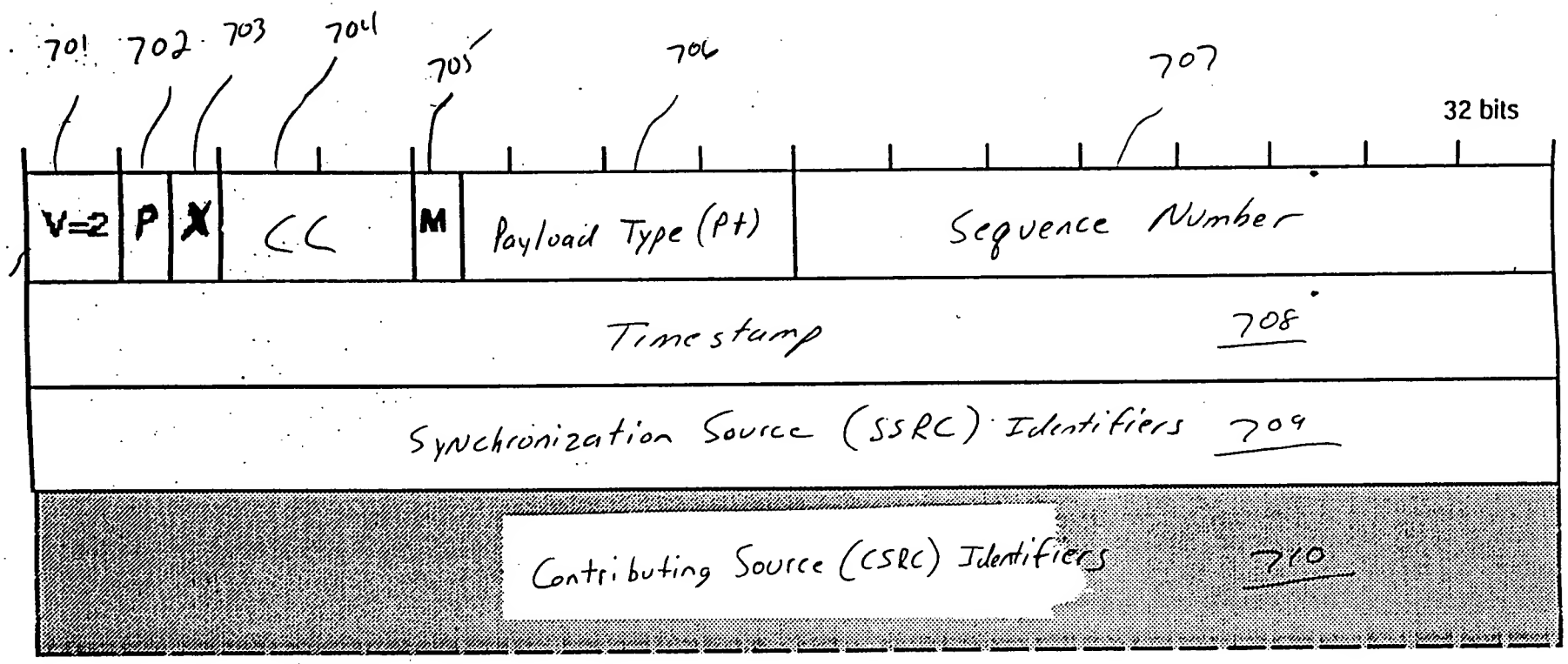

Fig. 7

(Prior Art)

8/18
YOR9-2000-0154-US1 (8728-376)

RTCCP

RTP initiated around RecoVC.xxx with X=1
or Changes of CODEC:

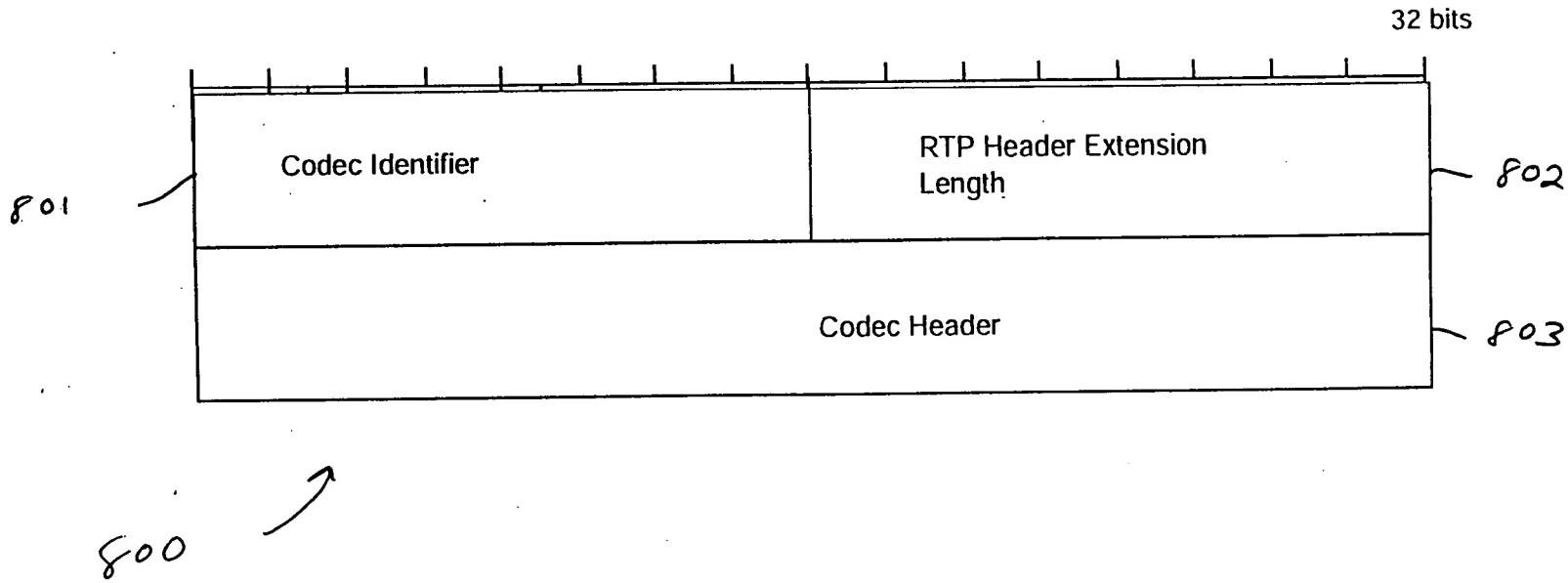

Fig. 8

DOCKET NO. 8728-376

9/18
YOR9-2000-0154US1 (8728-376)

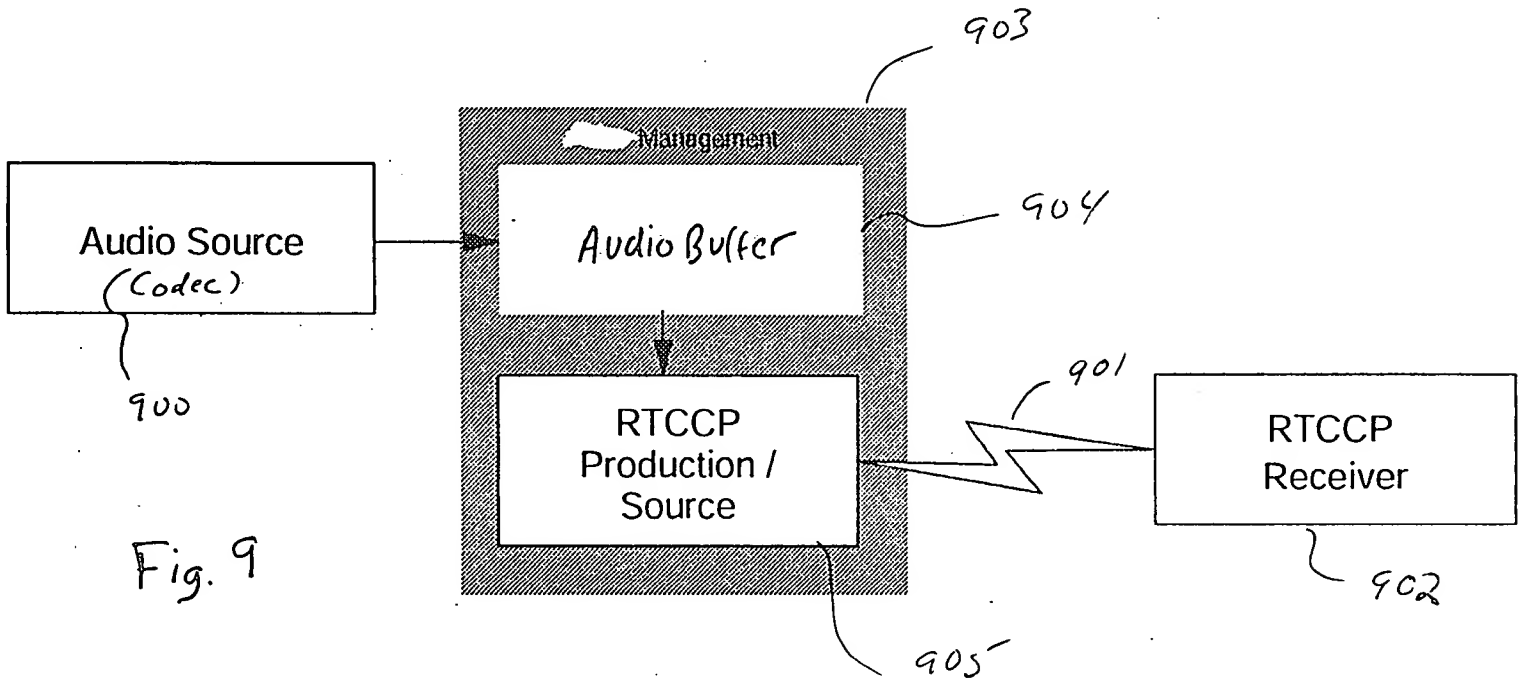

Fig. 9

Protocols structure:

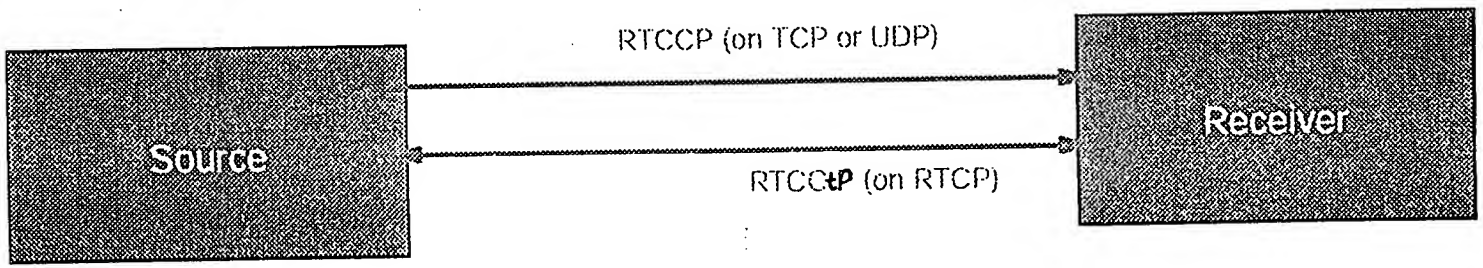

Fig. 11

Packet-Based format:

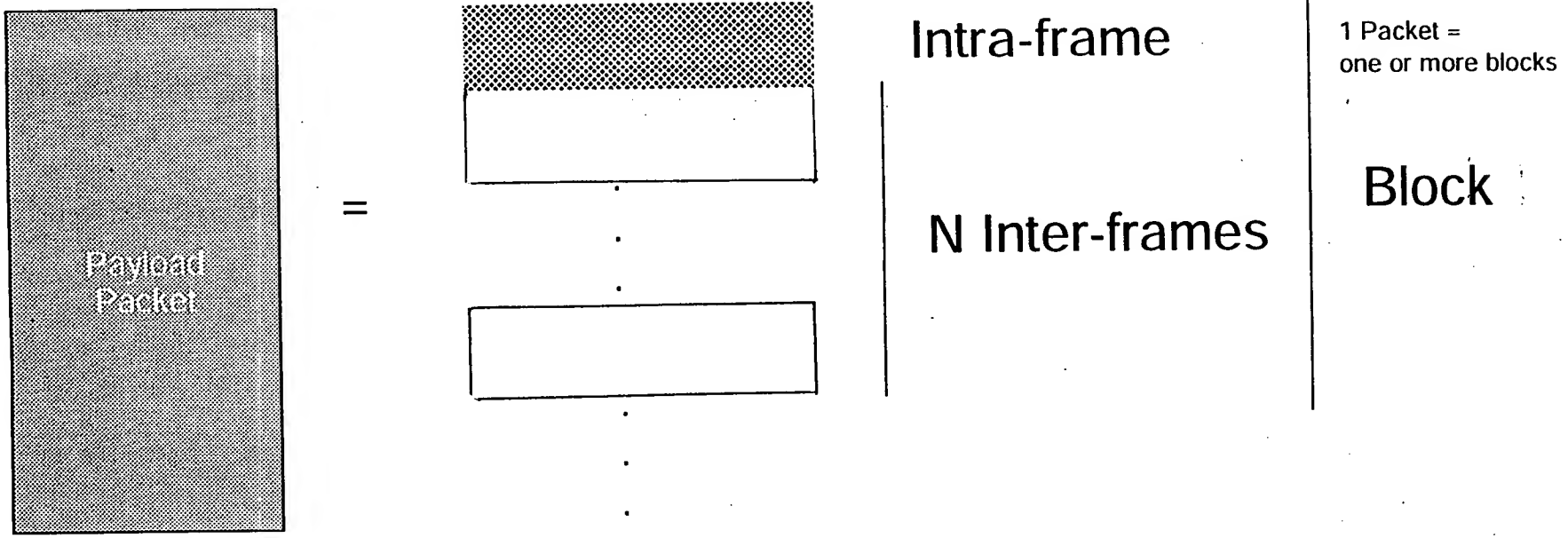

Fig. 10

10/18
VOB9-2000-015US1 (8128-326)

DRAFT 11/18

YOR9-2000-0054US1 (8208-376)

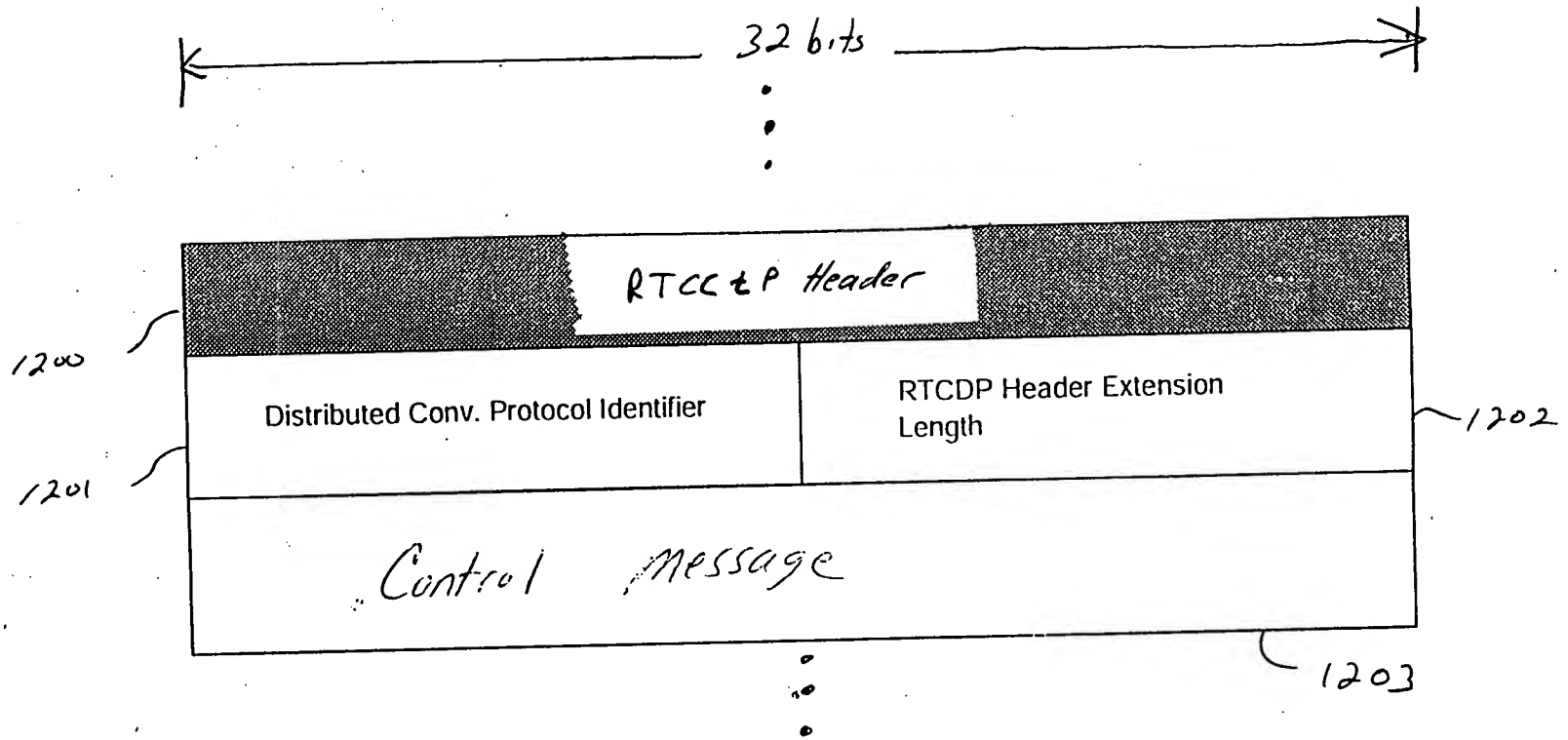

Fig. 12

DRAFT "nc5e0260

RTCDP wrapper/proxy for browsers and servers

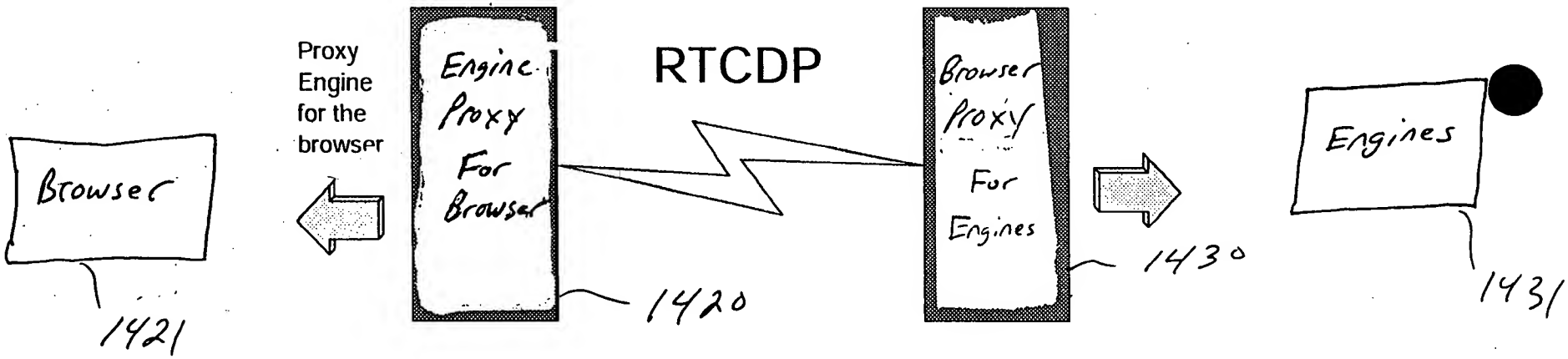

Fig. 14a

12/18
YOR9-2000-0154 US1 (8728-376)

Protocols structure:

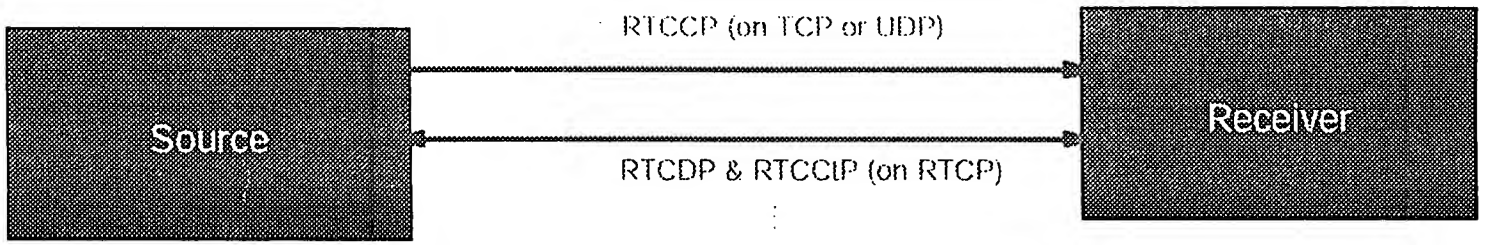

Fig. 15

GOTOFF "4252260

13/18
YOR9-2000-0154US (8128-326)

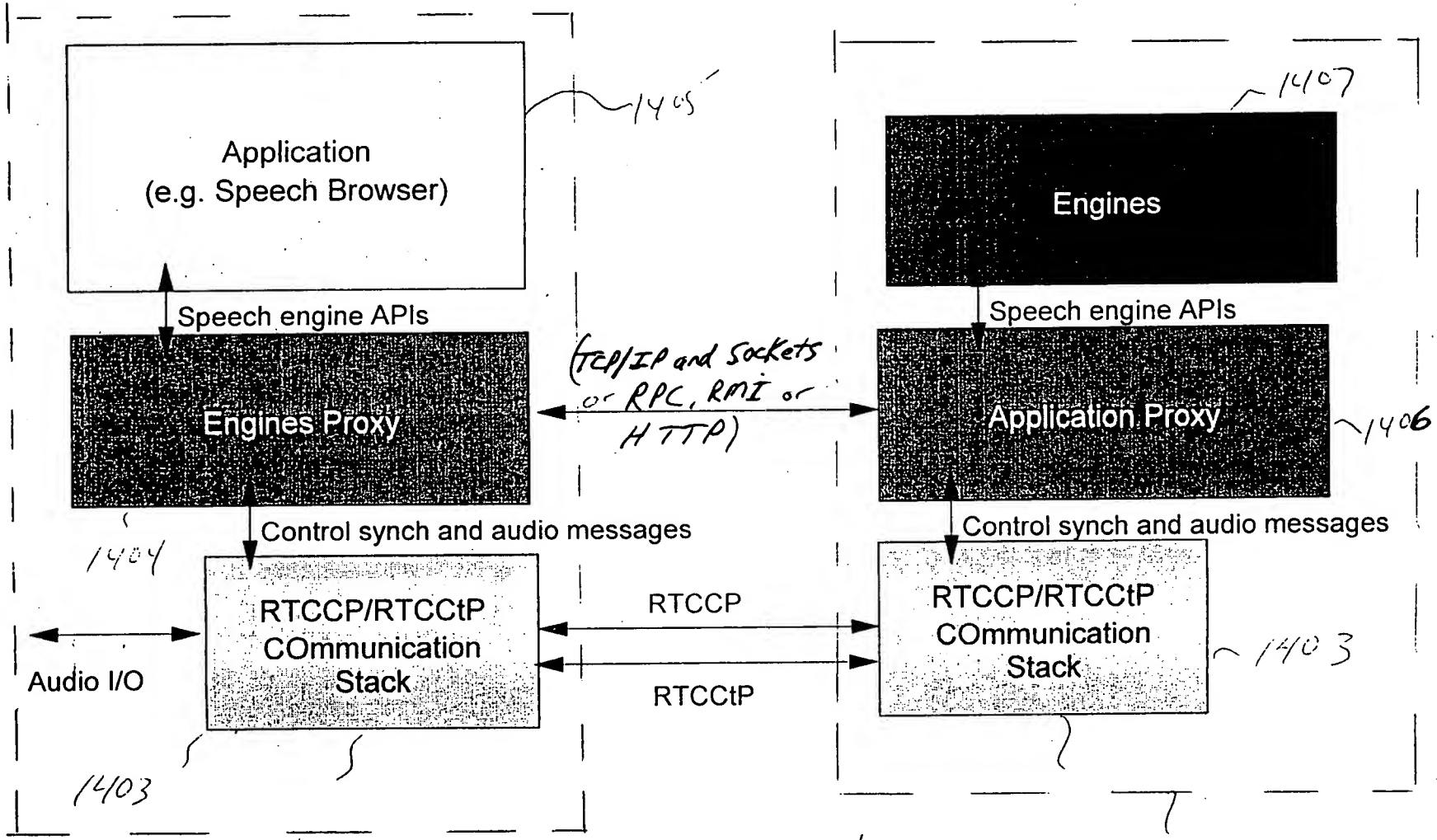

Fig. 14/6

09T0TT"h2SE0260

Conversational protocols (Umbrella stack)

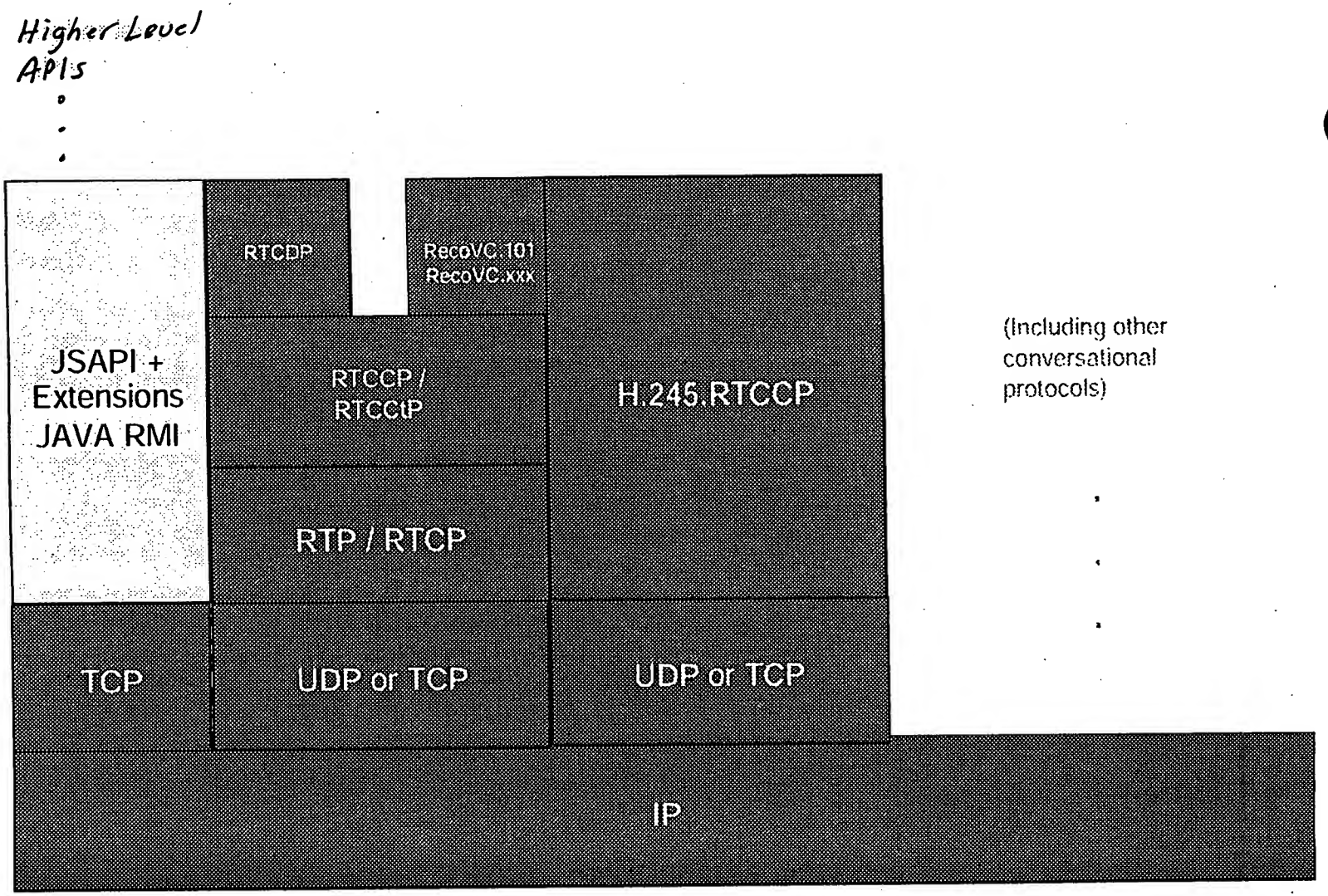

Fig. 15

14/18
YOR9-2000-0154US1 (8228-370)

0070FF 1259260

15/18
YOR9-2000-0154US1 (8726-376)

RTSCDP - Real Time Streaming Conversational Distributed Protocols

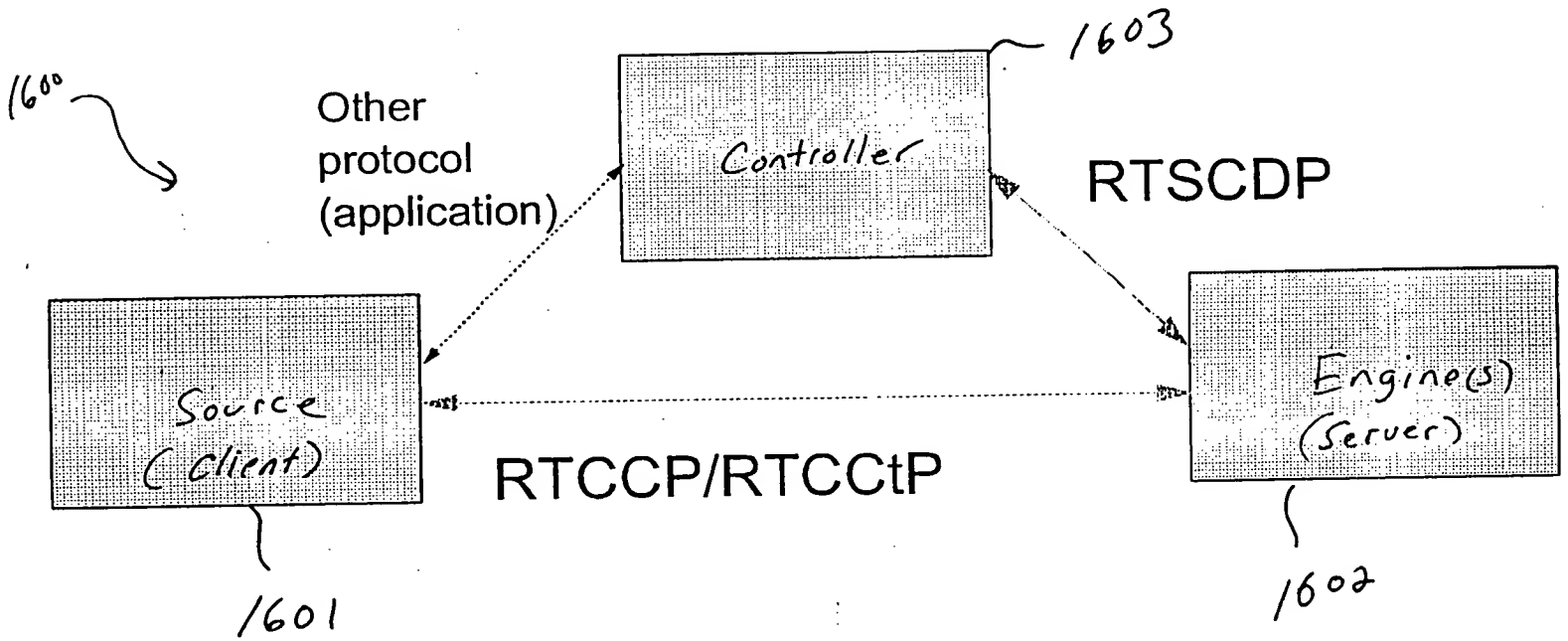

Fig. 16

YOR9-2000-0154US1 (8228-376)
16/18

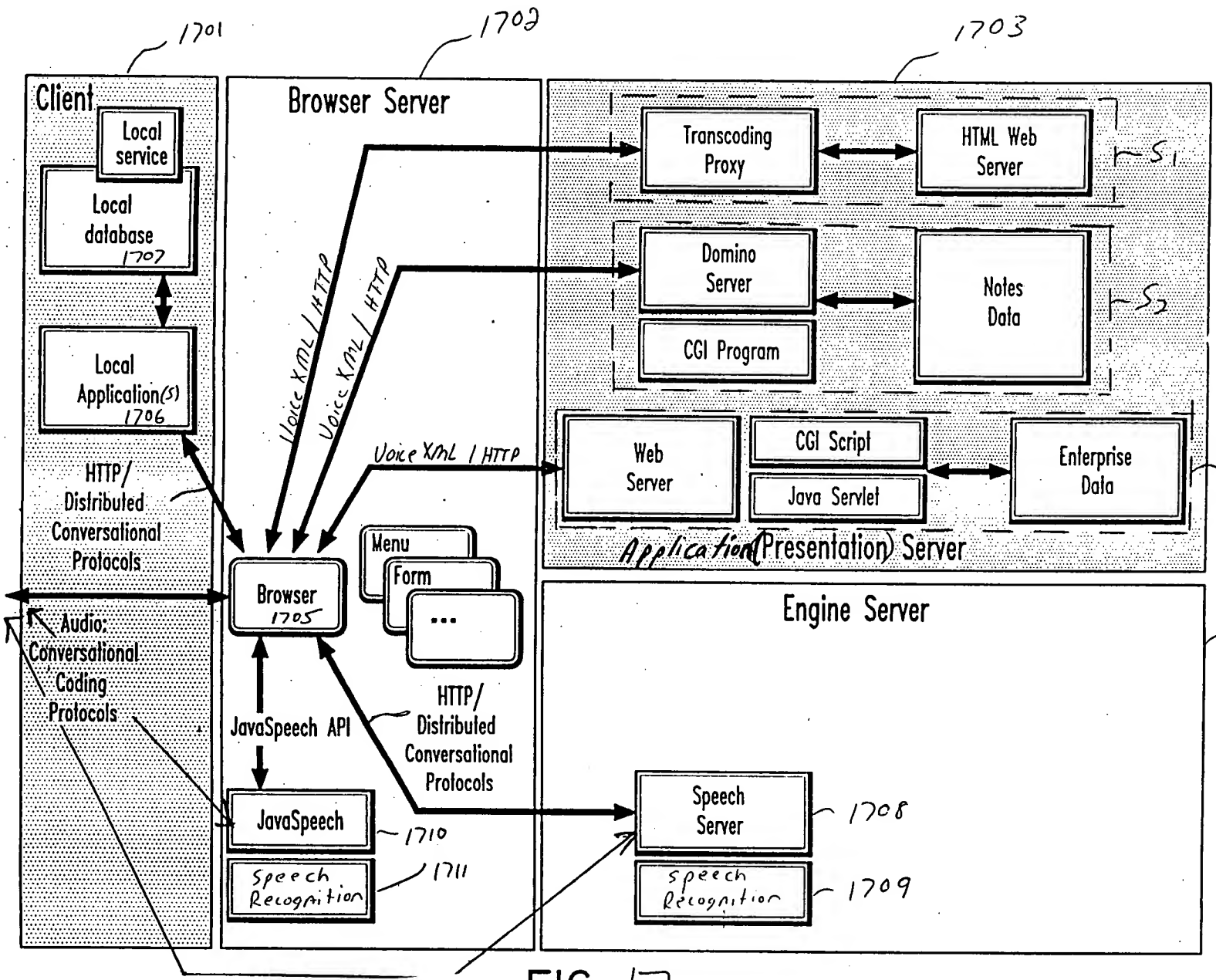

FIG. 17

17/18
YOR9-2000-0154US1 (8728-326)

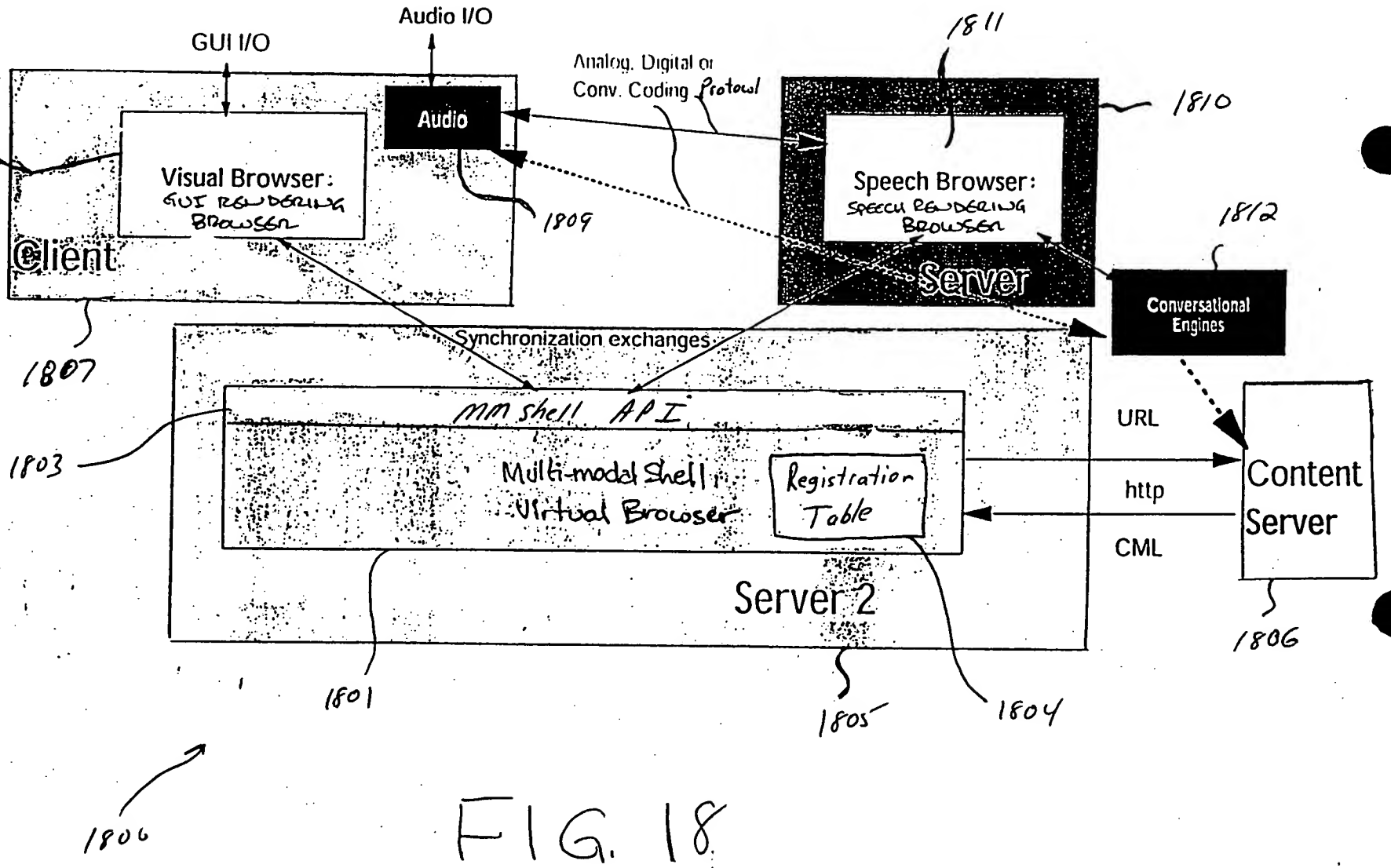

FIG. 18

Conversational Voice Communications

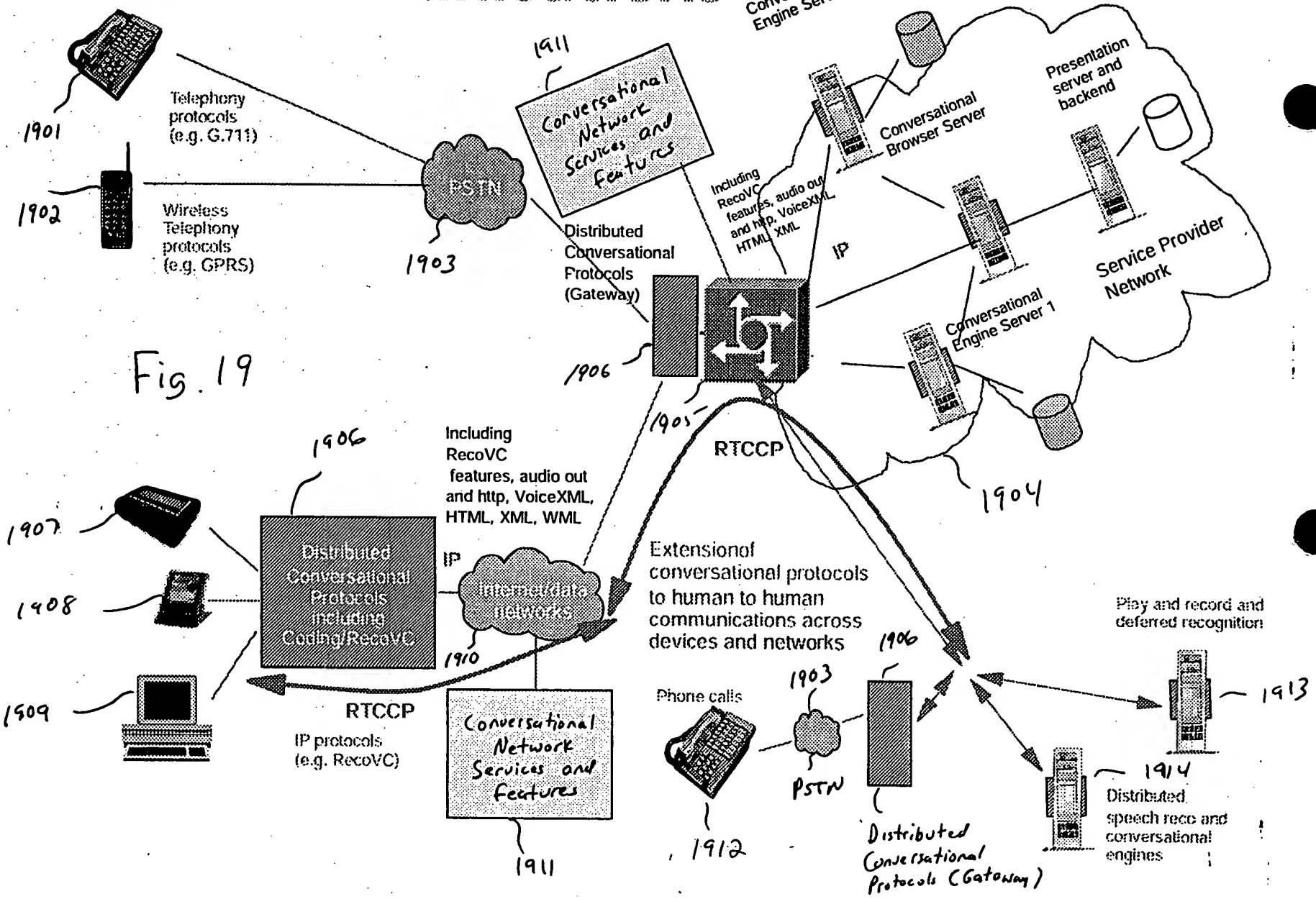

Fig. 19

18/18
YOR9-2000-0154US1 (8228-376)